Available online at www.sciencedirect.com

SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER

Linear Algebra and its Applications 404 (2005) 27–57

www.elsevier.com/locate/laa

Pole cancellation

Harry Dym [a,*], Shahar Nevo [b]

[a] *Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel*
[b] *Department of Mathematics, Bar Ilan University, Ramat Gan 52900, Israel*

Received 22 April 2004; accepted 16 February 2005
Available online 9 April 2005
Submitted by L. Rodman

To Peter Lancaster, valued friend and colleague, on the occasion of his 75th birthday

Abstract

The problem of eliminating the right half plane poles of an rmvf (rational matrix valued function) $G(z)$ with minimal realization $G(z) = D + C(zI_n - A)^{-1}B$ by multiplication on the right by a suitably chosen J-inner rmvf $\Theta(z)$ is considered from a number of different points of view, including the notion of minimal J conjugators that was introduced by Kimura, the null/pole structure of rmvf's that is developed at length in the monograph of Ball–Gohberg–Rodman, and the theory of reproducing kernel Hilbert spaces. Connections between these different points of view are developed and correspondences between (1) the Jordan chains corresponding to the right half plane eigenvalues of A^*, (2) the left null chains of $\Theta(z)$ in the sense of Ball–Gohberg–Rodman, and (3) the invariant subspaces of the generalized backwards shift operator applied to a suitably defined space of rmvf's are established. Enroute, a theorem of Kimura that relates the existence of minimal pole conjugators to the existence of solutions of a related Riccati equation is refined and made more precise with the aid of the techniques referred to above.

AMS classification: 30C15; 30E05; 47A68; 93B55; 93D99

Keywords: Pole cancellation; J-lossless conjugators; J-inner matrix valued functions; Riccati equations; Reproducing kernels; Smith–McMillan forms; Stability

* Corresponding author.
E-mail addresses: harry.dym@weizmann.ac.il (H. Dym), nevosh@math.biu.ac.il (S. Nevo).

doi:10.1016/j.laa.2005.02.014

1. Introduction

In this paper we deal with the problem of stabilizing an rmvf (rational matrix valued function) $G(z)$ in the open right half plane Π_+, by multiplying it on the right by an rmvf $\Theta(z)$ that is J-inner w.r.t. Π_+, where J is a given constant signature matrix. This problem arises in the theory of H_∞-control and has already been treated from various points of view. There are points of contact with [13,14], and, as the reviewer kindly pointed out [5,12]. The starting point of this investigation was the following theorem, which corresponds to Theorem 5.2 in the book of Kimura [18], in which

$$\#(\sigma(A) \cap \Omega) = \text{the number of eigenvalues of } A \text{ in } \Omega, \text{ counting multiplicities}$$

and $\mathrm{Mcdeg}(F)$ denotes the McMillan degree of a rmvf $F(z)$:

Theorem 1.1. *Let $G(z)$ be a proper rmvf with minimal realization*

$$G(z) = C(zI_n - A)^{-1}B + D, \tag{1.1}$$

where $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. Then there exists a proper rmvf $\Theta(z)$ that is J-inner w.r.t. Π_+ such that

(i) *$G(z)\Theta(z)$ is holomorphic in Π_+ and*
(ii) *$\mathrm{Mcdeg}(\Theta) = \#(\sigma(A) \cap \Pi_+)$,*

if and only if the Riccati equation

$$XA + A^*X - XBJB^*X = 0 \tag{1.2}$$

has a solution $X \geqslant 0$ such that

$$\widehat{A} = A - BJB^*X \quad \text{is stable}. \tag{1.3}$$

In this case,

$$\Theta(z) = -JB^*(zI_n + A^*)^{-1}XB + I_m = -JB^*X(zI_n - \widehat{A})^{-1}B + I_m \tag{1.4}$$

up to a right constant J-unitary multiplier. Moreover, if

$$H(z) = \widetilde{C}(zI_n - A)^{-1}B + \widetilde{D} \tag{1.5}$$

for some matrices $\widetilde{C}$, $\widetilde{D}$ of suitable sizes (where $(\widetilde{C}, A)$ is not necessarily observable), then

$$H(z)\Theta(z) = (\widetilde{C} - \widetilde{D}JB^*X)(zI_n - \widehat{A})^{-1}B + \widetilde{D}. \tag{1.6}$$

The minimality of the realization (1.1) is crucial. Indeed, if, for example, we take $C = 0$, then any J-inner mvf $\Theta(z)$ (w.r.t. Π_+) that is analytic in $\overline{\Pi_+}$ with $\mathrm{Mcdeg}(\Theta) = \#(\sigma(A) \cap \Pi_+)$ meets the conditions (i) and (ii) whether or not (1.2) has a solution $X \geqslant 0$ that satisfies (1.3). The proof of Kimura however, though essentially correct, seems to disregard the central role of the minimality of the realization (1.1). Kimura's proof is based on methods drawn from H_∞-control theory.

We shall establish Theorem 1.1 by a different approach that relies on the definition of the poles and zeros of a rmvf by the local Smith–McMillan form. In addition, we use the theory of pole structure and null-structure for rmvf's (see [3] for square rmvf's and [19] for general rmvf's). Using these tools we get a simpler proof.

The analogous problem of cancelling the zeros of a rmvf by multiplying on the right by a *J*-inner rmvf is treated in [10].

Another approach, based on the theory of reproducing kernel Krein spaces and reproducing kernel Hilbert spaces (RKKS, RKHS) and a theorem of de Branges was used in [7].

Before presenting the main result of this approach, we need some background material.

Let $(A, B) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times m}$ be a controllable pair, let

$$F(z) = B^*(zI_n + A^*)^{-1} \tag{1.7}$$

and let X be an $n \times n$ Hermitian matrix. Then the space

$$\mathscr{M}_X = \{F(z)Xu : u \in \mathbb{C}^n\} \tag{1.8}$$

with indefinite inner product

$$\langle FXu, FXv\rangle_{M_X} = v^*Xu$$

is a RKKS; it is an RKHS if and only if $X \geqslant 0$. The reproducing kernel (RK) is

$$K_\omega(z) = F(z)XF(\omega)^*.$$

If K_ω can be expressed as

$$K_\omega(z) = \frac{J - \Theta(z)J\Theta(\omega)^*}{z + \overline{\omega}},$$

where $\Theta(z)$ is an $m \times m$ rmvf, then $\mathscr{M}_X$ is said to be a dBK space $\mathscr{K}(\Theta)$ if X is Hermitian, and a de Branges space $\mathscr{H}(\Theta)$ if $X \geqslant 0$.

Theorem 4.1 of [7] (adapted to the current notation) states that:

Theorem 1.2. *Let (A, B) be a controllable pair and let $F(z)$ be as in (1.7). Then the RKKS $\mathscr{M}_X$ is a dBK space $\mathscr{K}(\Theta)$ if and only if the Hermitian matrix X is a solution of the Riccati equation*

$$A^*X + XA - XBJB^*X = 0. \tag{1.9}$$

In this case the rmvf $\Theta(z) = \Theta_X(z)$ is uniquely determined by the formula

$$\Theta_X(z) = -B^*(zI_n + A^*)^{-1}XBJ + I_m$$

up to a J-unitary constant multiplier on the right, and the following two identities hold:

$$X(zI_n - A)^{-1}BJ\Theta_X(z) = (zI_n + A^*)^{-1}XBJ,$$
$$(zI_n - A)^{-1}BJ\Theta_X(z) = (zI_n - \widehat{A})^{-1}BJ,$$

where $\widehat{A} = A - BJB^*X$.

The rmvf $\Theta_X(z)$ is J-inner if and only if $X \geqslant 0$. It can be expressed in terms of $\widehat{A}$ as

$$\Theta_X(z) = -B^*X(zI_n - \widehat{A})^{-1}BJ + I_m \tag{1.10}$$

up to a J-unitary constant multiplier on the right.

Theorem 1.2 generalizes the direction ($\Leftarrow$) of Theorem 1.1. The rmvf $\Theta_X(z)$ from Theorem 1.2 is related to the rmvf $\Theta(z)$ from Theorem 1.1 by the formula $J\Theta_X(z)J = \Theta(z)$ (the value at infinity here is taken to be I_m).

The paper is organized as follows: Section 2 is devoted to preliminaries. Section 3 presents the definition of a minimal pole conjugator w.r.t. Π_+ for rmvf's and a basic lemma. In Section 4.1, Theorem 1.1 is reformulated (as Theorem 4.4) to exhibit the fact that the minimal conjugator depends only on A and B. The role of the assumption that (A, B) is controllable is discussed in Section 4.2. Section 5 deals with the problem of cancelling poles of a general rmvf that is not necessarily proper. We note that finding a minimal pole conjugator w.r.t. Π_- for a rmvf $G(z)$, i.e., the anti-stabilizing case, can be treated completely analogously. However, we treat only the stabilizing case. In Section 6 we study the space $\mathscr{M}_X$ of (1.8) under the action of the backward shift operator

$$R_\alpha(f)(z) = \frac{f(z) - f(\alpha)}{z - \alpha}, \quad \alpha \in \mathbb{C}, \tag{1.11}$$

when the matrix X is the (unique) positive semidefinite solution of (1.2) that satisfies (1.3). Finally, in Section 7, we discuss some connections between the left null chains of $\Theta(z)$ in the sense of [3] and the Jordan chains of the matrix A^*.

The following theorem estabishes the equivalence of a number of different concepts that are discussed in this paper. A proof will be furnished in Section 6.

Theorem 1.3. *Let $G(z)$ be a rmvf with minimal realization*

$$G(z) = C(zI_n - A))^{-1}B + D,$$

where $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$ and $(A, B) \sim \left(\begin{bmatrix} A_+ & 0 \\ 0 & A^- \end{bmatrix}, \begin{bmatrix} B_+ \\ B_- \end{bmatrix}\right)$ (written conformally); $\sigma(A_+) \subset \Pi_+$, $\sigma(A_-) \subset \Pi_-$. Let X be an $n \times n$ Hermitian matrix. Then the following conditions are equivalent:

(1) *The matrix X is a positive semi-definite solution of (1.2), such that $\widehat{A} = A - BJB^*X$ is stable.*
(2) *There exists a minimal pole conjugator $\Theta(z)$ of $G(z)$ w.r.t. Π_+.*
(3) *There exists a minimal pole conjugator $\Theta(z)$ of the pair (A, B) w.r.t. Π_+.*
(4) *There exists a J-inner rmvf $\Theta(z)$ that is analytic in a neighborhood of $\overline{\Pi_+}$ such that the pair (A_+, B_+) is a left null pair of $\Theta(z)$ w.r.t. Π_+.*
(5) *The space $(\mathscr{M}_X, \langle , \rangle_{\mathscr{M}_X})$ is an $\mathscr{H}(\Theta)$ space for a J-inner rmvf of class $H_\infty^{m\times m}(\Pi_+)$, and*

$$\frac{\det \Theta(z)}{\det(zI_n - A)} \in H_\infty(\Pi_+). \tag{1.12}$$

The problem of constructing a minimal pole conjugator $\Theta(z)$ for a given rmvf $G(z)$ can also be studied from the point of view of interpolation theory. To illustrate this connection, suppose, for example, that $\Theta(z)$ is a rmvf that is J-inner w.r.t. Π_+ such that

$$\begin{bmatrix} \xi_j^* & \eta_j^* \end{bmatrix} \Theta(z_j) = 0$$

for a set of vectors $\xi_1, \ldots, \xi_k \in \mathbb{C}^p$, $\eta_1, \ldots, \eta_k \in \mathbb{C}^q$ and points $z_1, \ldots, z_k \in \Pi_+$. Then the linear fractional transformation

$$\begin{aligned}(\Theta_{11}\varepsilon + \Theta_{12})(\Theta_{21}\varepsilon + \Theta_{22})^{-1} \\ = \Theta_{12}\Theta_{22}^{-1} + (\Theta_{11} + \Theta_{12}\Theta_{22}^{-1}\Theta_{21})\varepsilon(\Theta_{21}\varepsilon + \Theta_{22})^{-1}\end{aligned}$$

based on the block decomposition of Θ that is conformal with J, maps mvf's ε that belong to the Schur class of $p \times q$ mvf's that are analytic and contractive in Π_+ into mvf's $s(z)$ of the same class that meet the interpolation conditions $\xi_j^* s(z_j) = \eta_j^*$ for $j = 1, \ldots, k$. Indeed, it was basically this strategy that was pursued recursively in [4]. We shall not enter into these issues in this paper, since there is an extensive literature on interpolation theory, see e.g., [3,6,11,17] and the references cited therein. The papers [8,9] exploit the spaces $\mathscr{M}_X$ based on solutions $X \geqslant 0$ of a Riccati equation analogous to (1.2) to study interpolation problems.

Analogous pole cancellation problems that are formulated for the open unit disc instead of the open half plane Π_+ will be considered in a separate publication.

2. Preliminaries

2.1. Notation

An mvf (matrix valued function) is said to be a meromorphic mvf if all its entries are meromorphic; it is said to be a rational mvf (rmvf) if all its entries are rational. The local behavior of a rmvf and of a meromorphic mvf obey the same rules; however, we treat only rmvf's because all the matrices that we deal with in this paper are such. The (i, j) entry of a matrix $A(z), B(z), \ldots, \widehat{A}(z), \widehat{B}(z), \ldots$ will be denoted by $a_{ij}(z), b_{ij}(z), \ldots, \widehat{a}_{ij}(z), \widehat{b}_{ij}(z), \ldots$

1. The *normal rank* of a rmvf $G(z)$ is the order of the largest square submatrix of $G(z)$ that is invertible except for at most a finite number of points. If the normal rank of a $p \times m$ rmvf $G(z)$ is equal to $\min\{p, m\}$, then G is said to be of full-rank; if in addition $p = m$, then G is said to be *regular*.
2. A rmvf $G(z)$ is called proper or analytic at infinity if $\lim_{z\to\infty} G(z) = G(\infty)$ is finite. If, in addition $G(\infty)$ is invertible, then $G(z)$ is said to be invertible at infinity.

2.2. The local Smith–McMillan form, zeros and poles

1. Let $G(z)$ be a $p \times m$ rmvf with normal rank r and let $z_0 \in \mathbb{C}$. Then

$$G(z) = E(z)\Lambda(z)F(z),$$

where $E(z)$, $F(z)$ are regular, analytic and invertible at z_0,

$$\Lambda(z) = \begin{bmatrix} D(z) & O_{r\times(m-r)} \\ O_{(p-r)\times r} & O_{(p-r)\times(m-r)} \end{bmatrix},$$

$D(z) = \mathrm{diag}\{(z-z_0)^{k_1}, (z-z_0)^{k_2}, \ldots (z-z_0)^{k_r}\}$ and

$$k_1 \leqslant k_2 \leqslant \cdots \leqslant k_r \tag{2.1}$$

are integers. The rmvf $\Lambda(z)$ is called the local Smith–McMillan form (local SM-form) of $G(z)$ at z_0, and is unique. The k_i's are referred to as the indices of the local SM-form of $G(z)$ at z_0. Observe that post-multiplying or pre-multiplying $G(z)$ by a rmvf that is analytic and invertible at z_0 does not affect the local SM-form.

2. A point $z_0 \in \mathbb{C}$ is called a zero (pole) of a $p \times m$ rmvf $G(z)$ if at least one of the indices k_i in (2.1) is positive (negative). In this case the set $\{k_i : k_i > 0\}$ ($\{-k_i : k_i < 0\}$) are the zero (pole) multiplicities of G at z_0. The total zero (pole) multiplicity of $G(z)$ at z_0 is equal to

$$M_\zeta(G; z_0) = \sum_{j=1}^{r} \max\{k_j, 0\} \left(M_\pi(G; z_0) = \sum_{j=1}^{r} \max\{-k_j, 0\} \right).$$

If $\Omega \subset \mathbb{C}$, then the total pole (zero) multiplicity of $G(z)$ w.r.t. Ω is

$$M_\pi(G; \Omega) = \sum_{z\in\Omega} M_\pi(G; z) \quad \left(M_\zeta(G; \Omega) = \sum_{z\in\Omega} M_\zeta(G; z) \right).$$

Remark 2.1. Poles are easy to locate, since z_0 is a pole of G if and only if z_0 is a pole of some entry of G. If $p = m$ and G is regular, then the multiplicities of z_0 as a zero of $G(z)$ coincide with the multiplicities of z_0 as a pole of $G(z)^{-1}$ and vice versa. In particular, $M_\zeta(G; z_0) = M_\pi(G^{-1}; z_0)$.

2.3. J-inner rmvf's

Let J be an $m \times m$ signature matrix, i.e., $J = J^* = J^{-1}$. A square rmvf $\Theta(z)$ is said to be J-unitary on $\mathrm{i}\mathbb{R}$ if $\Theta^*(z)J\Theta(z) = J$ at every point $z \in \mathrm{i}\mathbb{R}$ at which $\Theta(z)$ is analytic. In this case, by analytic continuation,

$$\Theta^\#(z)J\Theta(z) = J \quad \text{on } \mathbb{C},$$

where

$$\Theta^{\#}(z) = \Theta(-\overline{z})^*.$$

If, in addition,

$$\Theta^*(z) J \Theta(z) \leqslant J \quad (\Theta^*(z) J \Theta(z) \geqslant J)$$

for every point of analyticity of Θ in Π_+, then Θ is said to be J-inner ($-J$-inner) w.r.t. Π_+. Observe that Θ is J-unitary on $i\mathbb{R}$ if and only if Θ^{-1} is such, and Θ is J-inner w.r.t. Π_+ if and only if Θ^{-1} is $-J$-inner w.r.t. Π_+.

2.4. Stability and similarity

We assume a basic knowledge of realization theory, but, for the convenience of the reader, we shall review some known theorems and definitions.

Theorem 2.2. *The pair* $(C, A) \in \mathbb{C}^{p\times n} \times \mathbb{C}^{n\times n}$ *is observable if and only if for each nonzero vector* $v \in \mathbb{C}^n$ *the rmvf* $C(zI_n - A)^{-1}v$ *has a pole. The pair* $(A, B) \in \mathbb{C}^{n\times n} \times C^{n\times m}$ *is controllable if and only if for each nonzero vector* $u \in \mathbb{C}^n$ *the rmvf* $u^*(zI - A)^{-1}B$ *has a pole.*

Definition 2.3. An rmvf $G(z)$ is called *stable* if $M_\pi(G, \overline{\Pi_+}) = 0$.

Theorem 2.4. *Let* $G(z) = C(zI_n - A)^{-1}B + D$.

(a) *If* $z_0 \in \mathbb{C}$ *is a pole of* $G(z)$, *then* $z_0 \in \sigma(A)$. *The converse is valid if the realization is minimal.*
(b) *If* D *is invertible (and hence* G *is square) and* z_0 *is a zero of* G, *then* $z_0 \in \sigma(A - BD^{-1}C)$. *In the other direction, if the realization is minimal and* D *is invertible and* $z_0 \in \sigma(A - BD^{-1}C)$, *then* z_0 *is a zero of* $G(z)$.

Definition 2.5. An $n \times n$ matrix A is stable (anti-stable) if $\sigma(A) \subset \Pi_-(\sigma(A) \subset \Pi_+)$. The pair (A, B) is stabilizable (anti-stabilizable) if there is a matrix F such that $A + BF$ is stable (anti-stable).

We recall the well-known characterizations of controllability and stabilizability.

Lemma 2.6 (PBH-Test)

(a) *The pair* $(A, B) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times m}$ *is controllable if and only if*

$$\operatorname{rank}[\lambda I_n - A \;\; B] = n \quad \textit{for every point } \lambda \in \mathbb{C}.$$

(b) *The pair* (A, B) *is stabilizable if and only if*

$$\operatorname{rank}[\lambda I_n - A \;\; B] = n \quad \textit{for every point } \lambda \in \overline{\Pi_+}.$$

Corollary 2.7. *If (A, B) is controllable, then (A, B) is stabilizable. The converse holds if $\sigma(A) \subset \overline{\Pi_+}$.*

Definition 2.8. Let $(A_i, B_i, C_i) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times m} \times \mathbb{C}^{p\times n}$ for $i = 1, 2$. Then:

(a) The triple (A_1, B_1, C_1) is similar to the triple (A_2, B_2, C_2) if there exists an invertible matrix T such that $T^{-1}A_1T = A_2$, $T^{-1}B_1 = B_2$, $C_1T = C_2$.
(b) The pair (A_1, B_1) is left similar to the pair (A_2, B_2) if there exists an invertible matrix T such that $T^{-1}A_1T = A_2$, $T^{-1}B_1 = B_2$.
(c) The pair (C_1, A_1) is right similar to the pair (C_2, A_2) if there exists an invertible matrix T such that $T^{-1}A_1T = A_2$, $C_1T = C_2$.

Since we will not treat right similarity, both similarity and left similarity will be denoted by $\sim$.

Remark 2.9. Similarity is an RST relation in all of the three cases considered in the previous definition. If $(A_1, B_1, C_1) \sim (A_2, B_2, C_2)$, then $C_1(zI_n - A_1)^{-1}B_1 = C_2(zI_n - A_2)^{-1}B_2$. If (C_i, A_i) is observable and (A_i, B_i) is controllable for $i = 1, 2$, then the similarity matrix T is unique (in all cases, i.e., similarity, right similarity, and left similarity). Also, observability, controllability and the stabilizing property (see Definition 2.5) are preserved under similarity.

Theorem 2.10. *The pair (A, B) is not controllable if and only if*

$$(A, B) \sim \left(\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \begin{bmatrix} B_1 \\ 0 \end{bmatrix}\right) \quad \textit{and} \quad size(A_{11}) < size(A),$$

where A_{11}, A_{22} are square matrices and the blocks of the right pair are written conformally.

Remark 2.11. If $B \neq 0$, then (A_{11}, B_1) may be assumed to be controllable, as follows by applying Theorem 2.10 sequentially.

Theorem 2.12. *Let $G(z)$ be a proper rmvf with two minimal realizations:*

$$G(z) = C_1(zI_n - A_1)^{-1}B_1 + D = C_2(zI_n - A_2)^{-1}B_2 + D.$$

Then $(A_1, B_1, C_1) \sim (A_2, B_2, C_1)$ and the similarity matrix T connecting these two realizations is unique.

Definition 2.13. Let (A, B) be a controllable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. Suppose that $(A, B) \sim \left(\begin{bmatrix} A_+ & 0 \\ 0 & A- \end{bmatrix}, \begin{bmatrix} B_+ \\ B_- \end{bmatrix}\right)$, where $\sigma(A_+) \subset \Pi_+$ and $\sigma(A_-) \subset \Pi_-$. Then (A_-, B_-) and (A_+, B_+) will be referred to as *stable* and *anti-stable* parts of (A, B), respectively.

Stable and anti-stable parts of (A, B) are defined up to left similarity, i.e., if

$$(A, B) \sim \left(\begin{bmatrix} A_+^i & 0 \\ 0 & A_-^i \end{bmatrix}, \begin{bmatrix} B_+^i \\ B_-^i \end{bmatrix} \right), \quad i = 1, 2, \ \ \sigma(A_+^i) \subset \Pi_+, \ \ \sigma(A_-^i) \subset \Pi_-,$$

then there exists an invertible matrix S such that

$$S^{-1} \begin{bmatrix} A_+^1 & 0 \\ 0 & A_-^1 \end{bmatrix} S = \begin{bmatrix} A_+^2 & 0 \\ 0 & A_-^2 \end{bmatrix}, \quad S^{-1} \begin{bmatrix} B_+^1 \\ B_-^1 \end{bmatrix} = \begin{bmatrix} B_+^2 \\ B_-^2 \end{bmatrix},$$

and hence

$$\begin{aligned} & S^{-1} \begin{bmatrix} (zI - A_+^1)^{-1} & 0 \\ 0 & (zI - A_-^1)^{-1} \end{bmatrix} \begin{bmatrix} B_+^1 \\ B_-^1 \end{bmatrix} \\ &\quad = \begin{bmatrix} (zI - A_+^2)^{-1} & 0 \\ 0 & (zI - A_-^2)^{-1} \end{bmatrix} \begin{bmatrix} B_+^2 \\ B_-^2 \end{bmatrix}. \end{aligned}$$

Moreover, since $\left(\begin{bmatrix} A_+^{(i)} & 0 \\ 0 & A_-^{(i)} \end{bmatrix}, \begin{bmatrix} B_+^{(i)} \\ B_-^{(i)} \end{bmatrix} \right)$ is a controllable pair, $(A_+^{(i)}, B_+^{(i)})$ and $(A_-^{(i)}, B_-^{(i)})$ are controllable pairs for $i = 1, 2$. Thus, upon writing $S^{-1} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$ conformally, we obtain

$$(zI_k - A_+^2)^{-1} B_+^2 - S_{11}(zI_k - A_+^1)^{-1} B_+^1 = S_{12}(zI_\ell - A_-^1)^{-1} B_-^1.$$

Therefore, $S_{12}(zI_\ell - A_-^1)^{-1} B_-^1 \equiv 0$, which in turn implies that $S_{12} = 0$. Thus, (I_k, A_+^2, B_+^2) and (S_{11}, A_+^1, B_+^1) are two minimal realizations of the same rmvf. Consequently, they are similar. In particular, $(A_+^1, B_+^1) \sim (A_+^2, B_+^2)$. In the other direction, it is easy to prove that if $(A, B) \sim \left(\begin{bmatrix} A_+ & 0 \\ 0 & A_- \end{bmatrix}, \begin{bmatrix} B_+ \\ B_- \end{bmatrix} \right)$ and $(\widehat{A}_+, \widehat{B}_+) \sim (A_+, B_+)$, then $(A, B) \sim \left(\begin{bmatrix} \widehat{A}_+ & 0 \\ 0 & A_- \end{bmatrix}, \begin{bmatrix} \widehat{B}_+ \\ B_- \end{bmatrix} \right)$. Similarly, stable parts are only defined up to left similarity.

2.5. Kalman's Theorem and pole structure theory

It is now appropriate to introduce a basic notion from pole structure theory that is taken from [3].

Definition 2.14. Let $G(z)$ be a $p \times m$ rmvf and let $\Omega \subset \mathbb{C}$. A controllable pair $(A, B) \in \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times m}$ is called a left pole pair of $G(z)$ w.r.t. Ω if

(i) $\sigma(A) \subset \Omega$ and
(ii) there is a $p \times n$ matrix C such that (C, A) is observable and $G(z) - C(zI_n - A)^{-1} B$ is analytic in some neighborhood of Ω.

The following theorem plays a central role in the problem of cancelling poles.

Theorem 2.15 [16]. *Let $G(z)$ be a proper rmvf with minimal realization $G(z) = C(zI_n - A)^{-1}B + D$. Then the order of A is equal to $M_\pi(G; \mathbb{C})$. Moreover, for every pole z_0 of $G(z)$,*

$$M_\pi(G; z_0) = \text{the algebraic multiplicity of } z_0 \text{ as an eigenvalue of } A.$$

This theorem justifies the use of the term $\mathrm{Mcdeg}(G)$ (which is equal $M_\pi(G; \mathbb{C})$) for the order of the matrix A in the minimal realization of G. A proof of this theorem can be found in [16]. Another proof can be based on the pole structure theory that is developed in [2,3].

Corollary 2.16. *In Definition 2.14*:

(1) $A \in \mathbb{C}^{n\times n}$, *where* $n = M_\pi(G; \Omega)$.
(2) *The matrix C is unique*.
(3) *The left pole pair* (A, B) *is determined up to left similarity*.

Corollary 2.17. *If $H(z)$ is proper and invertible at ∞, then $M_\pi(H; \mathbb{C}) = M_\zeta(H; \mathbb{C})$.*

Proof. This follows from Theorem 2.15 and from the inversion formula for realizations: If $G(z) = C(zI_n - A)^{-1}B + D$ and D is invertible, then

$$G(z)^{-1} = -D^{-1}C(zI_n - (A - BD^{-1}C))^{-1}BD^{-1} + D^{-1}. \qquad \square$$

3. Cancelling the poles of an rmvf in Π_+

We turn now to a definition that leads us to the essence of this paper. From now on J is a given fixed $m \times m$ signature matrix.

Definition 3.1. Let $G(z)$ be a $p \times m$ rmvf. A proper $m \times m$ rmvf $\Theta(z)$ that is J-inner w.r.t. Π_+ is called a *minimal pole conjugator* of $G(z)$ w.r.t. Π_+ if the following two conditions are in force:

(a) $M_\pi(G\Theta; \Pi_+) = 0$.
(b) $\mathrm{Mcdeg}(\Theta) = M_\pi(G; \Pi_+)$.

Remark 3.2. If Θ is a minimal pole conjugator of G, w.r.t. Π_+, then so is ΘU, where U is any contant J-unitary matrix. Moreover, if $G(z)$ has no poles on $\mathrm{i}\mathbb{R}$, then condition (a) in the preceding definition is equivalent to requiring that $G\Theta$ is stable.

Lemma 3.3. *Let $G(z)$ be a $p \times m$ rmvf with a pole at z_0, let $H(z)$ be an $m \times k$ rmvf of normal rank m (which forces $m \leqslant k$) and suppose further that $G(z)H(z)$ is holomorphic at z_0. Then z_0 is a zero of $H(z)$. Moreover, if the first ℓ indices in the local SM-form of $G(z)$ at z_0 are negative, i.e., if $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_\ell < 0$, then the last ℓ indices in the local SM-form of $H(z)$ are positive: $0 < j_{m-\ell+1} \leqslant j_{m-\ell+2} \leqslant \cdots \leqslant j_m$, and*

$$k_1 + j_m \geqslant 0, \quad k_2 + j_{m-1} \geqslant 0, \ldots, k_\ell + j_{m-\ell+1} \geqslant 0.$$

In particular,

$$M_\pi(G; z_0) \leqslant M_\zeta(H; z_0).$$

Proof. Assume first that $p \leqslant m$ and $m = k$. Then $G(z) = E(z)[M\ 0]F(z)$, where $F(z)$ and $E(z)$ are analytic and invertible at z_0,

$$M(z) = \operatorname{diag}\{(z - z_0)^{k_1}, \ldots, (z - z_0)^{k_\ell}, 0 \ldots, 0\}$$

is a $p \times p$ rmvf, $k_1 \leqslant \cdots \leqslant k_r$, r is the normal-rank of G, and $k_{\ell+1}, \ldots k_r$ are non-negative integers. Similarly,

$$H(z) = U(z)L(z)V(z), \quad L(z) = \begin{bmatrix} (z-z_0)^{j_1} & & 0 \\ & \ddots & \\ 0 & & (z-z_0)^{j_m} \end{bmatrix},$$

where $j_1 \leqslant \cdots \leqslant j_r$, $U(z)$ and $V(z)$ are holomorphic and invertible at z_0. Since GH is holomorphic at z_0, the rmvf $[M\ 0]FUL$ is also holomorphic at z_0. Moreover, since FU is analytic and invertible at z_0, there is a permutation $\{i_1, \ldots, i_m\}$ of $\{1, \ldots, m\}$ such that $(FU)(z_0)_{t,i_t} \neq 0$ for $1 \leqslant t \leqslant m$. Thus, $[M\ 0]FU$ has pole of order $-k_t$ in the (t, i_t) entry, for $1 \leqslant t \leqslant \ell$, and consequently, $j_{i_t} \geqslant -k_t$. Since $j_1 \leqslant \cdots \leqslant j_m$ and $k_1 \leqslant \cdots \leqslant k_r$, the assertion follows for $p \leqslant m$ and $m = k$.

The proof for the case $p \leqslant m$ and $m < k$ goes along the same lines, the only difference being that now it is necessary to analyze $[M\ O]FU[L\ O]$ instead of $[M\ O]FUL$. However, the zeros and poles of these two rmvf's are exactly the same.

The case $p > m$ is treated in much the same way. □

Remark 3.4. The lemma is false when $m > k$, even if H is of full-rank. For example, if

$$G(z) = \begin{bmatrix} 0 & \dfrac{1}{z - z_0} \end{bmatrix} \quad \text{and} \quad H(z) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

then GH is holomorphic at z_0, but the conclusions of the lemma are false.

Now if $\Theta(z)$ is any proper J-inner rmvf such that $G(z)\Theta(z)$ is holomorphic in Π_+, then, by Lemma 3.3,

$$M_\zeta(\Theta; \Pi_+) \geqslant M_\pi(G; \Pi_+).$$

Moreover, since

$$\mathrm{Mcdeg}\,(\Theta) = M_\pi(\Theta; \mathbb{C}) = M_\zeta(\Theta; \mathbb{C}) \geqslant M_\zeta(\Theta; \Pi_+),$$

it follows that

$$\mathrm{Mcdeg}(\Theta) \geqslant M_\pi(G; \Pi_+)$$

and hence that the McMillan degree of a proper J-inner rmvf $\Theta(z)$ that stabilizes $G(z)$ cannot be less than $M_\pi(G; \Pi_+)$. In other words, a minimal pole conjugator $\Theta(z)$ of $G(z)$ w.r.t. Π_+ has minimal possible McMillan degree. Therefore, the zeros of such a rmvf Θ are located exactly at the poles of G in Π_+ and have the same multiplicities. In addition, since $\Theta^\# J\Theta = J$, it follows that Θ is a regular rmvf and $\Theta^{-1} = J\Theta^\# J$. Therefore, λ is a pole (zero) of Θ if and only if $-\overline{\lambda}$ is a zero (pole) of Θ, and of the same multiplicities. Thus, Θ is analytic in a neighborhood of $\overline{\Pi_+}$.

Remark 3.5. The Riccati equation (1.2) does not admit a stabilizing solution when $\sigma(A) \cap \mathrm{i}\mathbb{R} \neq \emptyset$; see e.g., Lemma 3.6 in [18]. Moreover, multiplication by mvf's $\Theta(z)$ that are J-inner w.r.t. Π_+ is not an effective way of cancelling poles on the imaginary axis. Indeed, if $\lambda \in \mathrm{i}\mathbb{R}$ is a pole of $G(z)$ and $\Theta(z)$ is a J-inner mvf w.r.t. Π_+ such that $G(z)\Theta(z)$ is holomorphic at λ, then, by Lemma 3.3, λ is a zero of $\Theta(z)$ of the same multiplicity, and hence also a pole of $\Theta(z)$ of that multiplicity. Thus, $G(z)\Theta(z)$ will be holomorphic on $\mathrm{i}\mathbb{R}$ only if the poles on $\mathrm{i}\mathbb{R}$ contributed by $\Theta(z)$ are cancelled by $G(z)$. This will only happen in special cases. If, for example,

$$\omega = \mathrm{i} + \delta|a|^2 \quad \text{with } |a| = |b| \neq 0 \quad \text{and} \quad \delta > 0$$

and

$$G(z) = (z+1)^{-1}\begin{bmatrix} z - \mathrm{i} & \frac{z-\omega}{z-\mathrm{i}} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ \frac{\delta\overline{a}b}{z-\omega} & 1 \end{bmatrix},$$

then

$$\Theta(z) = I_2 - \frac{\delta v v^* J}{z - \mathrm{i}}, \quad \text{with } v = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

can be expressed as

$$\Theta(z) = \begin{bmatrix} z - \omega & 0 \\ -\delta\overline{a}b & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{z-\mathrm{i}} & 0 \\ 0 & z - \mathrm{i} \end{bmatrix}\begin{bmatrix} 1 & \frac{\delta a\overline{b}}{z-\omega} \\ 0 & \frac{1}{z-\omega} \end{bmatrix}.$$

Consequently,

$$G(z)\Theta(z) = (z+1)^{-1}\begin{bmatrix} z - \omega & 1 + \delta a\overline{b} \end{bmatrix}$$

is holomorphic on $\overline{\Pi_+}$.

4. Cancelling the poles of a proper rmvf

In this section we deal with the problem of finding the minimal pole conjugator of a proper rmvf, that has no poles on $\mathrm{i}\mathbb{R}$.

4.1. Background and formulation of the main theorem

Lemma 4.1. *Let* $(C, A) \in \mathbb{C}^{p\times n} \times \mathbb{C}^{n\times n}$ *be an observable pair and let* $u(z)$ *be an* $n \times 1$ *rmvf that is analytic at* $z = \alpha$. *Then* $(zI_n - A)^{-1}u(z)$ *has pole at* α, *if and only if* $C(zI_n - A)^{-1}u(z)$ *has a pole at* α.

Proof. It suffices to consider the case that A is in Jordan form and then, since each Jordan cell interacts with a different set of columns of C, to focus on the case of a single Jordan cell

$$A = \alpha I_n + N, \quad \text{where } N = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Then

$$\begin{aligned}(zI_n - A)^{-1}u(z) &= \left(\frac{I_n}{z-\alpha} + \frac{N}{(z-\alpha)^2} + \cdots + \frac{N^{n-1}}{(z-\alpha)^n}\right) \\ &\quad \cdot \left(u(\alpha) + (z-\alpha)u'(\alpha) + (z-\alpha)^2\frac{u''(\alpha)}{2!} + \cdots\right) \\ &= \frac{1}{z-\alpha}\left(u(\alpha) + Nu'(\alpha) + \cdots + N^{n-1}\frac{u^{(n-1)}(\alpha)}{(n-1)!}\right) \\ &\quad + \frac{1}{(z-\alpha)^2}\left(Nu(\alpha) + N^2u'(\alpha) + \cdots + N^{n-1}\frac{u^{(n-2)}(\alpha)}{(n-2)!}\right) \\ &\quad + \cdots + \frac{1}{(z-\alpha)^n}N^{n-1}u(\alpha) + g(z),\end{aligned}$$

where $g(z)$ is an $n \times 1$ rmvf that is holomorphic at $z = \alpha$. The last formula can be written more transparently in terms of the vector

$$v = u(\alpha) + Nu'(\alpha) + \cdots + N^{n-1}\frac{u^{(n-1)}(\alpha)}{(n-1)!}$$

as

$$(zI_n - A)^{-1}u(z) = \frac{v}{z-\alpha} + \frac{Nv}{(z-\alpha)^2} + \cdots + \frac{N^{n-1}v}{(z-\alpha)^n} + g(z).$$

Thus, $C(zI_n - A)^{-1}u(z)$ will be holomorphic at $z = \alpha$ if and only if

$$Cv = CNv = \cdots = CN^{n-1}v = 0.$$

Since (C, A) is observable if and only if (C, N) is observable, we see that $C(zI_n - A)^{-1}u(z)$ is holomorphic at $z = \alpha$ if and only if $v = 0$ (see Theorem 2.2), i.e., if and only if $(zI_n - A)^{-1}u(z)$ is holomorphic at $z = \alpha$.

To put it another way, we have shown that if $(zI_n - A)^{-1}u(z)$ has a pole at $z = \alpha$, then $C(zI_n - A)^{-1}u(z)$ has a pole at $z = \alpha$. It is clear that if $(zI_n - A)^{-1}u(z)$ is holomorphic at $z = \alpha$, then so is $C(zI_n - A)^{-1}u(z)$. $\square$

Lemma 4.1 is false if $u(z)$ is not holomorphic at the point $z = \alpha$. The choice $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $C = [1 \quad 0]$ and $u(z) = \begin{bmatrix} -z^{-1} \\ 1 \end{bmatrix}$ serves as an example.

Lemma 4.1 implies that the minimal pole conjugator $\Theta(z)$ w.r.t Π_+ of a rmvf $G(z)$ with minimal realization

$$G(z) = D + C(zI_n - A)^{-1}B$$

depends only upon the matrices A and B and not upon the matrices C and D. Moreover, since

$$M_\pi(G; \Pi_+) = \#(\sigma(A) \cap \Pi_+),$$

by Theorem 2.15, the definition of a minimal pole conjugator of $G(z)$ w.r.t. Π_+ can now be formulated totally in terms of the pair (A, B) as follows:

Definition 4.2. Let (A, B) be a controllable pair and assume that $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. A proper rmvf $\Theta(z)$ that is J-inner w.r.t. Π_+ is called a minimal pole conjugator of the pair (A, B) w.r.t. Π_+ if

(i) $(zI_n - A)^{-1}B\Theta(z)$ is stable;
(ii) $\mathrm{Mcdeg}(\Theta) = \#(\sigma(A) \cap \Pi_+)$, counting multiplicities.

Moreover, if (ii) is in force, then (i) holds if and only if

(i′) $\Theta(z)$ is a minimal pole conjugator of the rmvf $G(z) = C(zI_n - A)^{-1}B + D$ w.r.t. Π_+ for at least one (and hence every) choice of C and D such that the pair (C, A) is observable.

Remark 4.3. If $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$ and $T^{-1}AT = \begin{bmatrix} A_+ & 0 \\ 0 & A_- \end{bmatrix}$ and $T^{-1}B = \begin{bmatrix} B_+ \\ B_- \end{bmatrix}$ conformally, where $\sigma(A_+) \subset \Pi_+$ and $\sigma(A_-) \subset \Pi_-$, then, by Lemma 2.6, (A, B) is stabilizable if and only if (A_+, B_+) is controllable.

Corollary 4.4. *Let (A, B) be a controllable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$ and let $\Theta(z)$ be a minimal pole conjugator of the pair (A, B) w.r.t. Π_+. Then $\Theta(z)$ is analytic in a neighborhood of $\overline{\Pi_+}$.*

This corollary follows at once from the discussion following Remark 3.4.

Theorem 4.5. *Let $G(z)$ be an rmvf with minimal realization $G(z) = C(zI_n - A)^{-1} + D$, let (A_+, B_+) be the anti-stable part of the pair (A, B) and suppose that $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. Then the following are equivalent*:

(1) *$\Theta(z)$ is a minimal pole conjugator of $G(z)$ w.r.t. Π_+.*
(2) *$\Theta(z)$ is a minimal pole conjugator of the pair (A, B) w.r.t. Π_+.*
(3) *$\Theta(z)$ is a minimal pole conjugator of the pair (A_+, B_+) w.r.t. Π_+.*

Proof. The equivalence of (1) and (2) follows from the discussion in Definition 4.2. The equivalence of (2) and (3) follows from Definition 4.2 and Corollary 4.4. $\square$

Motivated by Corollary 2.7, Remark 4.3 and Theorem 4.5, we can now reformulate Theorem 1.1 as follows:

Theorem 4.6. *Let (A, B) be a stabilizable pair and assume that $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. Then there exists a minimal pole conjugator $\Theta(z)$ of the pair (A, B) w.r.t. Π_+ if and only if the Riccati equation*

$$XA + A^*X - XBJB^*X = 0 \tag{4.1}$$

has a solution $X \geqslant 0$ such that

$$\widehat{A} = A - BJB^*X \quad \text{is stable}. \tag{4.2}$$

In this case every such Θ is of the form

$$\Theta(z) = -JB^*(zI_n + A^*)^{-1}XB + I_m = -JB^*X(zI_n - \widehat{A})^{-1}B + I_m, \tag{4.3}$$

up to a constant J-unitary multiplier on the right, and if $G(z) = C(zI_n - A)^{-1}B + D$, then

$$G(z)\Theta(z) = (C - DJB^*X)(zI_n - \widehat{A})^{-1})B + D. \tag{4.4}$$

Left null pairs may be defined in the same terms that were used to define left pole pairs in Definition 2.14. However, for our purposes the following equivalent definition that is based on Theorem 1.54 of [3] is more convenient.

Definition 4.7. Let $H(z)$ be a rmvf and let $\Omega \subset \mathbb{C}$. Then the pair (A, B) is said to be a left null pair for $H(z)$ w.r.t. Ω if the following four conditions are met:

(1) (A, B) is controllable.
(2) $(zI_n - A)^{-1}BH(z)$ is holomorphic in Ω.
(3) $\sigma(A) \subset \Omega$.
(4) $\operatorname{order}(A) = M_\zeta(H; \Omega)$.

The analysis in Chapters 1 and 3 of [3] yields the following conclusions:

Theorem 4.8. *If $H(z)$ is a regular rmvf then the following are equivalent*:

(1) *The pair (A, B) is a left null pair for $H(z)$ w.r.t. Ω.*
(2) *The pair (A, B) is a left pole pair for $H(z)^{-1}$ w.r.t. Ω.*
(3) *The pair (A, B) is controllable, $\sigma(A) \subset \Omega$ and there exists a matrix C such that (C, A) is observable and $C(zI_n - A)^{-1}B - H(z)^{-1}$ is holomorphic in some neighborhood of Ω.*

We remark that the matrix C in (3) of Theorem 4.8 is unique. Also, analogously to Corollary 2.16, a left null pair is defined up to left similarity.

Corollary 4.9. *If $\Theta(z)$ is a minimal pole conjugator of the pair (A, B) w.r.t. Π_+, and if (A_+, B_+) is the anti-stable part of (A, B), then (A_+, B_+) is a left null pair of $\Theta(z)$ w.r.t. Π_+.*

Theorem 4.10. *Let m, n_1, n_2 be positive integers, and let (A_1, B_1), (A_2, B_2) be two controllable pairs, where for $i = 1, 2$, $(A_i, B_i) \in \mathbb{C}^{n_i \times n_i} \times \mathbb{C}^{n_i \times m}$ and $\sigma(A_i) \cap \mathrm{i}\mathbb{R} = \emptyset$. Suppose that for $i = 1, 2$, $\Theta_i(z)$ is a minimal pole conjugator of the pair (A_i, B_i) w.r.t. Π_+ such that $\Theta_i(\infty) = I_m$. Then $\Theta_1(z) = \Theta_2(z)$ if and only if the anti-stable parts of (A_1, B_1) and (A_2, B_2) are left similar.*

Proof. Suppose that $\Theta_1(z) = \Theta_2(z) = \Theta(z)$ and, for $i = 1, 2$, let (A_i^+, B_i^+) be the anti-stable part of (A_i, B_i). By Theorem 4.5, $\Theta(z)$ is a minimal pole conjugator w.r.t. Π_+ for both (A_1^+, B_1^+) and (A_2^+, B_2^+) and each of these two pairs is a left null pair of $\Theta(z)$ w.r.t. Π_+. Thus, they are left similar. The other direction follows from the proof of Theorem 4.6, which displays the fact that formula (4.3) is invariant under left similarity of the anti-stable part of (A, B). $\square$

Lemma 4.11. *Let $G(z)$ be a $p \times m$ rmvf of full-rank $r = \min(p, m)$. Let $\widehat{G}(z)$ be the new rmvf that is formed by multiplying one row (resp. column) by $z - z_0$ if $r = p$ (resp. $r = m$). Let $k_1 \leqslant \cdots \leqslant k_r$ be the indices of the local SM-form of G at z_0, and let $\widehat{k}_1 \leqslant \cdots \leqslant \widehat{k}_r$ be the indices of the local SM-form of $\widehat{G}$ at z_0. Then $k_i \leqslant \widehat{k}_i$ for $1 \leqslant i \leqslant r$, and*

$$\sum_{i=1}^{r} \widehat{k}_t = 1 + \sum_{i=1}^{r} k_t,$$

i.e., only one index is changed, and it increases by 1.

Proof. We prove only the case $p \leqslant m$. The case $p > m$ follows analogously or by passing to transposes.

The proof is by finite induction on p. First, for a non-identically vanishing scalar rational function $f(z)$ and for $z_0 \in \mathbb{C}$, let $n(f, z)$ denote the index of first nonzero entry in the Laurent expansion of f about z_0:

$$n(f, z_0) = \max\{\ell : f(z)(z - z_0)^{-\ell} \text{ is analytic at } z_0\},$$

i.e.,

$$n(f, z_0) = \begin{cases} k & \text{if } f \not\equiv 0 \text{ and } z_0 \text{ is a zero of order } k \text{ of } f, \\ -t & \text{if } f \not\equiv 0 \text{ and } z_0 \text{ is a pole of order } t \text{ of } f, \\ 0 & \text{if } f \not\equiv 0 \text{ and } z_0 \text{ is not a pole or a zero of } f, \\ \infty & \text{if } f \equiv 0. \end{cases}$$

For a $q \times s$ rmvf $F(z)$, define

$$N(F, z_0) = \min\{n(f_{ij}(z), z_0) : 1 \leqslant i \leqslant q, 1 \leqslant j \leqslant s\}.$$

For $p = 1$, it is obvious that the local SM-form of $\widehat{G}$ at z_0 is

$$[(z - z_0)^{N(G,z_0)+1}, 0, \ldots, 0].$$

Suppose we have established the lemma for $p = k$ for some $k < m$ and let $p = k + 1$. Assume we multiplied the ℓth row of G by $z - z_0$ to get $\widehat{G}$. Let (i, j) be an entry of $\widehat{G}$, for which $N(\widehat{G}, z_0)$ is achieved. If this minimum is achieved both in the ℓth row of $\widehat{G}$ as well as in some other row of $\widehat{G}$, we take $i = \ell$.

Starting with $\widehat{G}$, perform the operations

$$C_t \mapsto C_t - \frac{\widehat{g}_{it}(z)}{\widehat{g}_{ij}(z)} C_j, \quad 1 \leqslant t \leqslant m, \ \ t \neq j \tag{4.5}$$

sequentially, and then

$$R_s \mapsto R_s - \frac{\widehat{g}_{sj}(z)}{\widehat{g}_{ij}(z)} R_i, \quad 1 \leqslant s \leqslant k+1, \ \ s \neq i, \tag{4.6}$$

where C_n or R_n denote the nth row or the nth column, respectively, of the rmvf on which the operation (4.5) or (4.6) is performed.

Let $\widetilde{G}(z)$ denote the new matrix that is obtained from $\widehat{G}(z)$ after completing all the indicated column operations (4.5) and row operations (4.6). The ith row and the jth column of $\widetilde{G}(z)$ are zero, except for $\widetilde{g}_{ij}(z) = \widehat{g}_{ij}(z)$.

These operations are performed by post-multiplying or by pre-multiplying, respectively, by a matrix, which by the definition of $N(\widehat{G}, z_0)$ is analytic and invertible at z_0, and thus does not affect the local SM-form at z_0. Now, consider the submatrix created from $\widetilde{G}$ by eliminating its ith row and jth column, and apply the induction assumption to complete the proof. □

Remark 4.12. If $\Theta(z)$ is a minimal pole conjugator of a proper $p \times m$ rmvf $G(z)$ w.r.t. Π_+ and if Mcdeg $\Theta = n$, then Θ cancels n right half plane poles of G, but introduces n new left half plane poles in the product $G\Theta$. However, since some of these may be cancelled by left half plane zeros of G, the most that can be said in general is that $\mathrm{Mcdeg}(G\Theta) \leqslant \mathrm{Mcdeg}(G)$. If $p \geqslant m$ and the left half plane zeros of a full rank proper $p \times m$ G do not cancel any of the left half plane poles of Θ, then by Lemma 4.11 equality prevails. For example, if $C(zI - A)^{-1}B + D$ is a minimal realization for $G(z)$ and if D is invertible and $\sigma((A - BD^{-1}C)) \cap \sigma((-A^*) \cap \Pi_-) = \emptyset$, then $\mathrm{Mcdeg}(G\Theta) = \mathrm{Mcdeg}(G)$ and the realization (4.4) is minimal.

The following well-known facts on the solutions of Riccati equations will be useful:

Lemma 4.13 [18, p. 43]. *The Riccati equation* (4.1) *has at most one Hermitian solution X such that the matrix $\widehat{A} = A - BJB^*X$ is stable.*

The solution X referred to in Lemma 4.13 will be denoted $R_{st}(A, B)$, if it exists. The following result is easily verified by direct computation.

Lemma 4.14. *Let $T \in \mathbb{C}^{n\times n}$ be invertible. Then $X = R_{st}(A, B)$, if and only if $T^*XT = R_{st}(T^{-1}AT, T^{-1}B)$.*

4.2. Proof of the main theorem

Before proceeding to the proof, it is convenient to recall the following result:

Theorem 4.15. *Let $\Theta(z)$ be an $m \times m$ proper regular rmvf, and let $C(zI_n - A)^{-1}B + D$ be a minimal realization of $\Theta(z)$. Then Θ is J-unitary on $\mathrm{i}\mathbb{R}$ if and only if*

(1) $D^*JD = J$.
(2) *There is an invertible Hermitian matrix H, such that* (i) $HA + A^*H = -C^*JC$ *and* (ii) $B = -H^{-1}C^*JD$.

Moreover, there is only one such H, and, Θ is J-inner ($-J$-inner) w.r.t. Π_+ if and only if $H > 0$ ($H < 0$).

This is Theorem 2.1 of [1]; see also Theorem 4.5 of [18].

Proof of Theorem 4.6. Suppose first that A is unstable, and let

$$\Theta(z) = C_c(zI - A_c)^{-1}B_c + I_m$$

be a minimal pole conjugator of the pair (A, B) w.r.t. Π_+, with $\Theta(\infty) = I_m$. Then

$$\Theta(z)^{-1} = -C_c(zI - (A_c - B_cC_c))^{-1}B_c + I_m \quad \text{and}$$
$$\sigma(A_c - B_cC_c) \subset \Pi_+. \tag{4.7}$$

Moreover, $\Theta(z)^{-1}$ is $-J$-inner w.r.t. Π_+ and the indicated realization in (4.7) is minimal. The inclusion in (4.7) follows from Remarks 2.1 and 3.4 since the zeros of $\Theta(z)$ are all in Π_+. By Theorem 4.15, there is unique $P < 0$ such that

$$P(A_c - B_cC_c) + (A_c^* - C_c^*B_c^*)P = -C_c^*JC_c \tag{4.8}$$

and

$$-JC_c = -B_c^*P. \tag{4.9}$$

Next, since $\sigma(A) \subset \Pi_+$, (A, B) is controllable. Thus, $\Theta(z)$ is a minimal pole conjugator of both the pair (A, B) and the rmvf $\Theta(z)^{-1}$. Moreover, as the matrices A and $A_c - B_cC_c$ are both anti-stable, Theorem 4.10 guarantees the existence of an invertible matrix T such that

$$T^{-1}B_c = B, \quad T^{-1}(A_c - B_cC_c)T = A. \tag{4.10}$$

Consequently, $C_c = JB^*T^*P$, and $B_c = TB$, by (4.9) and (4.10), and hence, (4.8) becomes

$$(-T^*PT)A + A^*(-T^*PT) = (-T^*PT)BJB^*(-T^*PT),$$

and so $X = -T^*PT > 0$ solves (4.1). Moreover, since A_c is stable, the matrix

$$\begin{aligned}\widehat{A} = A - BJB^*X &= T^{-1}(A_c - B_cC_c)T - T^{-1}B_cJB_c^*T^{-*}(-T^*PT)\\ &= T^{-1}A_cT\end{aligned}$$

is also stable. By (4.9) and (4.10), we have

$$\Theta(z)^{-1} = -JB^*X(zI_n - A)^{-1}B + I_m,$$

and thus,

$$\Theta(z) = -JB^*X(zI_n - \widehat{A})^{-1}B + I_m.$$

The second formula for $\Theta(z)$ in (4.3) is also deduced easily.

Now if $G(z) = C(zI_n - A)^{-1}B + D$, then

$$\begin{aligned}G(z)\Theta(z) &= C(zI_n - A)^{-1}B + C(zI_n - A)^{-1}BJB^*X(zI_n - \widehat{A})^{-1}B\\ &\quad -DJB^*X(zI_n - \widehat{A})^{-1}B + D\\ &= C(zI_n - A)^{-1}[(zI_n - \widehat{A}) + BJB^*X](zI_n - \widehat{A})^{-1}B\\ &\quad -DJB^*X(zI_n - \widehat{A})^{-1}B + D\\ &= C(zI_n - \widehat{A})^{-1}B - DJB^*X(zI_n - \widehat{A})^{-1}B + D,\end{aligned}$$

which justifies (4.4).

Next, we drop the assumption that A is unstable. Then there is an invertible S such that

$$S^{-1}AS = \begin{bmatrix} A_+ & 0 \\ 0 & A_- \end{bmatrix} \quad \text{and} \quad S^{-1}B = \begin{bmatrix} B_+ \\ B_- \end{bmatrix} \quad \text{(conformally)}, \tag{4.11}$$

where $\sigma(A_+) \subset \Pi_+$ and $\sigma(A_-) \subset \Pi_-$. By Remark 4.3, the pair (A_+, B_+) is controllable. Therefore, by Theorem 4.5, $\Theta(z)$ is a minimal pole conjugator of the pair

(A_+, B_+). Thus, by the first part of the proof, $X_0 = R_{st}(A_+, B_+)$, $X_0 > 0$ exists. Set

$$\widehat{X}_0 = \begin{bmatrix} X_0 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{4.12}$$

Then $\widehat{X}_0 = R_{st}\left(\begin{bmatrix} A_+ & 0 \\ 0 & A_- \end{bmatrix}, \begin{bmatrix} B_+ \\ B_- \end{bmatrix}\right)$, and hence, by Lemma 4.14,

$$X = S^{-*}\widehat{X}_0 S^{-1} = R_{st}(A, B) \quad \text{and} \quad X \geqslant 0. \tag{4.13}$$

Moreover, by (4.11) and (4.12), (4.3) and (4.4) hold, too. This completes the proof in one direction (⇒).

Next, to obtain the converse (⇐), suppose that the Riccati equation (4.1) admits a positive semidefinite stabilizing solution X, i.e., that $X = R_{st}(A, B)$ exists and $X \geqslant 0$, and let

$$\Theta(z) = -JB^*(zI_n + A^*)^{-1}XB + I_m. \tag{4.14}$$

Then, in view of (4.1),

$$(zI_n + A^*)^{-1}X = X(zI_n - \widehat{A})^{-1}. \tag{4.15}$$

Thus, $\Theta(z)$ can also be rewritten as

$$\Theta(z) = -JB^*X(zI_n - \widehat{A})^{-1}B + I_m, \tag{4.16}$$

and a straightforward calculation yields the formula

$$J - \Theta(\omega)^* J\Theta(z) = B^*(\overline{\omega}I_n - \widehat{A}^*)^{-1}X(z + \overline{\omega})(zI_n - \widehat{A})^{-1}B$$

at points of analyticity $z, w \in \mathbb{C}$ of Θ.

If $\omega = z$, then the expression on the right is equal to zero if $z \in i\mathbb{R}$ and is positive semidefinite if $z \in \Pi_+$. Therefore, Θ is J-inner w.r.t. Π_+. Moreover, formula (4.14) shows that if λ is a pole of $\Theta(z)$, then $-\overline{\lambda} \in \sigma(A)$. However, since this realization is not minimal (unless X is invertible), the most that one can say is that

$$M_\pi(\Theta; \lambda) \leqslant \text{algebraic multiplicity of } -\overline{\lambda} \text{ as an eigenvalue of } A.$$

Formula (4.16) shows that Θ is stable. Thus,

$$\text{Mcdeg}(\Theta) = M_\pi(\Theta; \mathbb{C}) = M_\pi(\Theta; \Pi_-) \leqslant \#(\sigma(A \cap \Pi_+))$$

(counting multiplicities). However, by Lemma 3.3,

$$\#(\sigma(A \cap \Pi_+)) = M_\pi(G; \Pi_+) \leqslant M_\zeta(\Theta; \Pi_+) = M_\pi(\Theta; \Pi_-) = \text{Mcdeg}(\Theta).$$

Therefore, Θ is a minimal pole conjugator of the pair (A, B) w.r.t. Π_+. □

Remark 4.16. Note that (4.3) determines Θ uniquely up to a multiplicative J-unitary constant $D_c = \Theta(\infty)$ on the right, since, by Lemma 4.13, there is at most one X equal to $R_{st}(A, B)$.

5. Cancelling the poles of a general rmvf

The objective of this section is find a necessary and sufficient condition for the existence of a minimal pole conjugator of a rmvf $G(z)$ w.r.t. Π_+ when $G(z)$ is not restricted to be proper. This will be achieved by reducing the problem to the case of a proper rmvf and then invoking Theorem 4.6.

The reduction of $G(z)$ from non proper to proper is attained by one of the following two methods:

(I) Discard the polynomial part of $G(z)$, i.e., since G is rational, there exists a matrix polynomial $P(z)$ such that $G_1(z) = G(z) - P(z)$ is proper.
or
(II) Set $G_2(z) = \frac{1}{(z+1)^k} G(z)$, where k is the maximal order of infinity as a pole of any entry of $G(z)$.

It is clear that the rmvf's $G(z)$ and $G_2(z)$ have the same pole multiplicities at every point $z_0 \in \overline{\Pi_+}$, since $-1 \in \Pi_-$. The same holds true for $G(z)$ and $G_1(z)$, since the addition of a rmvf that is analytic at z_0 does not change the pole multiplicities at z_0. Consequently, by applying Theorem 4.6 to either G_1 or G_2, we obtain the following conclusions for G:

Theorem 5.1. *Let $G(z)$ be a $p \times m$ non proper rmvf without poles on $\mathrm{i}\mathbb{R}$ and let $G_i(z) = C_i(zI_{n_i} - A_i)^{-1}B_i + D_i$, $i = 1, 2$, be the proper rmvf's that are constructed from G as above and suppose that the indicated realizations are minimal. Then (A_1, B_1) and (A_2, B_2) have similar anti-stable parts and there exists a minimal pole conjugator Θ of G w.r.t. Π_+ if and only if there exists a positive semi-definite matrix $X_i = R_{st}(A_i, B_i)$. In that case every such Θ is of the form*

$$\Theta(z) = -JB_i^*(zI_{n_i} + A_i^*)^{-1}X_iB_i + I_m = -JB_i^*X_i(zI_{n_i} - \widehat{A}_i)^{-1}B_i + I_m,$$

up to a constant J-unitary factor on the right, where $\widehat{A}_i = A_i - B_iJB_i^*X_i$ is stable.

Proof. Let Θ be a J-inner rmvf w.r.t. Π_+ that is holomorphic in $\overline{\Pi_+}$. Then, in view of the preceding discussion, $G\Theta$ has a pole $z_0 \in \overline{\Pi_+}$ if and only if $G_i\Theta$ has a pole at z_0. Furthermore, by the uniqueness of Θ with $\Theta(\infty) = I$, Theorem 4.10 implies that (A_1, B_1) and (A_2, B_2) have similar anti-stable parts and that the Riccati equations for $i = 1$ and for $i = 2$ are essentially the same. (Only the anti-stable part of (A_i, B_i) is significant for $R_{st}(A_i, B_i)$, see the proof of Theorem 4.6.) The last conclusion can also be deduced from the fact that, by Definition 2.14, the anti-stable part of (A_i, B_i) is a left pole pair of G_i over Π_+, and, by Corollary 2.16, these two anti-stable parts are left similar. □

6. R_α-invariance

In this section we study the behavior of the space

$$\mathscr{M}_X = \{F(z)Xu : u \in \mathbb{C}^n\}$$

(see (1.7), under the action of the backward shift operator R_α defined by formula (1.11).

We start with the following lemma.

Lemma 6.1. *Let (A, B) be a controllable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$ and assume that X is a Hermitian solution to (1.2). Then the space*

$$\mathscr{M}_X = \{F(\lambda)Xu : u \in \mathbb{C}^n\}$$

is invariant under the backwards shift operator R_α

$$R_\alpha : f \in \mathscr{M}_X \to \frac{f(\lambda) - f(\alpha)}{\lambda - \alpha}$$

for every point $\alpha \in \mathbb{C} \setminus \sigma(-A^*)$. Moreover, for every such point α, R_α is a one-to-one map of $\mathscr{M}_X$ onto $\mathscr{M}_X$ that is similar to multiplication by the matrix $-(\alpha I_n + A^*)^{-1}$.

Proof. First take $\alpha \in \mathbb{C} \setminus (\sigma(-A^*) \cup \sigma(\widehat{A}))$, $\widehat{A} = A - BJB^*X$. In this case the asserted invariance follows from the formulas

$$\begin{aligned}\frac{F(\lambda) - F(\alpha)}{\lambda - \alpha} &= B^* \left\{ \frac{(\lambda I_n + A^*)^{-1} - (\alpha I_n + A^*)^{-1}}{\lambda - \alpha} \right\} \\ &= -F(\lambda)(\alpha I_n + A^*)^{-1}\end{aligned} \tag{6.1}$$

and

$$(\alpha I_n + A^*)X = X(\alpha I_n - \widehat{A}).$$

Next, since the pair (A, B) is controllable, these formulas imply that if $f(\lambda) = F(\lambda)Xu$, then

$$(R_\alpha f)(\lambda) = 0 \Rightarrow (\alpha I_n + A^*)^{-1}Xu = 0 \Rightarrow Xu = 0 \Rightarrow f(\lambda) \equiv 0,$$

i.e., R_α is an injective map of $\mathscr{M}_X$ into itself. Therefore, since $\mathscr{M}_X$ is finite dimensional, R_α maps $\mathscr{M}_X$ onto $\mathscr{M}_X$.

Let T denote the operator from $\mathscr{M}_X$ onto $\mathscr{R}(X)$, the range of X, that is defined by the rule $T(F(z)Xu) = Xu$. T is well-defined since (A, B) is controllable, and it is a bijective map. Moreover,

$$TR_\alpha FXu = -(\alpha I_n + A^*)^{-1}Xu = -(\alpha I_n + A^*)^{-1}TFXu,$$

i.e.,

$$TR_\alpha = -(\alpha I_n + A^*)^{-1}T \quad \text{or} \quad TR_\alpha T^{-1} = -(\alpha I_n + A^*)^{-1}.$$

Suppose next that $\alpha \in (\mathbb{C} \setminus \sigma(-A^*))$, and let $\{\alpha_k\}$ be a sequence of points in $\mathbb{C} \setminus (\sigma(-A^*) \cup \sigma(\widehat{A}))$ that tend to α as $k \uparrow \infty$. Then, as

$$(R_\alpha F)(z)Xu = -F(z)(\alpha I_n + A^*)^{-1}Xu$$

and

$$(\alpha I_n + A^*)^{-1}Xu = \lim(\alpha_k I_n + A^*)^{-1}Xu$$

is the limit of a sequence of vectors that belong to the finite dimensional space $\mathscr{R}(X)$, the range of X, it follows that

$$(\alpha I_n + A^*)^{-1}Xu \in \mathscr{R}(X).$$

Consequently, $(R_\alpha F)(z)Xu$ belongs to $\mathscr{M}_X$ for such α also. □

Observe that Lemma 6.1 holds also under the assumptions that $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$, (A, B) is stabilizable and $X = R_{st}(A, B)$.

Lemma 6.2. *Let (A, B) be a stabilizable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$ and $X = R_{st}(A, B)$. Let $\lambda_1, \ldots, \lambda_k$ denote the distinct eigenvalues of A^* in Π_+ and let $\overline{\mu}_1, \ldots, \overline{\mu}_\ell$ denote the distinct eigenvalues of A^* in Π_-. Then*:

(1) *$\mathscr{R}(X)$, the range of X, is related to the generalized eigenspaces $V_\lambda(A^*)$ corresponding to the eigenvalues λ of A^* by the formula*

$$\mathscr{R}(X) = V_{\lambda_1}(A^*) \dot{+} \cdots \dot{+} V_{\lambda_k}(A^*). \tag{6.2}$$

(2) *$\mathscr{N}(X)$, the nullspace of X, is given by $\mathscr{N}(X) = V_{\mu_1}(A) \dot{+} \cdots \dot{+} V_{\mu_\ell}(A)$.*

Proof. Let the invertible matrices $S \in \mathbb{C}^{n\times n}$ and $X_0 \in \mathbb{C}^{\ell\times\ell}$ be defined by formulas (4.11)–(4.13). Then, since X_0 is invertible,

$$\mathscr{R}(X) = \left\{ S^{-*} \begin{bmatrix} u \\ 0 \end{bmatrix} : u \in \mathbb{C}^\ell \right\}. \tag{6.3}$$

Let

$$C_\lambda = \lambda I_k + N$$

be an upper triangular $k \times k$ Jordan cell in the Jordan form of A^* corresponding to a point $\lambda \in \sigma(A^*) \cap \Pi_+$, and let the columns of the $n \times k$ matrix W_λ be the corresponding Jordan chain. Then

$$A^* W_\lambda = W_\lambda C_\lambda \tag{6.4}$$

and hence, by formula (4.11),

$$\begin{bmatrix} A_+^* & 0 \\ 0 & A_-^* \end{bmatrix} S^* W_\lambda = S^* W_\lambda C_\lambda.$$

Thus, upon writing

$$S^*W_\lambda = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \tag{6.5}$$

conformally with the indicated block decomposition of A^*, we obtain the pair of formulas

$$A_+^* U_1 = U_1 C_\lambda \quad \text{and} \quad A_-^* U_2 = U_2 C_\lambda.$$

Moreover, since $A_-^* - \lambda I$ is invertible for $\lambda \in \Pi_+$, the latter implies that $U_2 = 0$, since

$$(A_-^* - \lambda I)^j U_2 = U_2 N^j \quad \text{for } j = 1, 2, \ldots$$

and $N^k = 0$. Therefore, by formulas (6.3) and (6.5),

$$\mathscr{R}(W_\lambda) \subset \mathscr{R}(X)$$

for every Jordan chain that corresponds to a Jordan cell C_λ with $\lambda \in \sigma(A^*) \cap \Pi_+$. Moreover, since the columns of the W_λ corresponding to distinct cells C_λ in the Jordan form of A^* are linearly independent and since A_+ and X_0 are both $\ell \times \ell$ matrices, the asserted result follows.

By analogous arguments it may be shown that $V_{\mu_1} \dot{+} \cdots \dot{+} V_{\mu_\ell}(A) \subset \mathscr{N}(X)$. Since $\mathscr{N}(X) \dot{+} \mathscr{R}(X) = \mathbb{C}^n$ and the sum of the algebraic multiplicities of $\lambda_1, \ldots, \lambda_k$ as eigenvalues of A^* and of $\mu_1, \ldots, \mu_\ell$ as eigenvalues of A is also equal to n, it follows that (2) holds, too. □

The dimensions of the spaces $\mathscr{R}(X)$ and $\mathscr{N}(X)$ may also be found in Lemma 3.6 of [18].

Lemma 6.3. *Let A be an $n \times n$ matrix and let $\omega \in \sigma(A^*)$ and V_ω be an $n \times k$ matrix such that*

$$A^* V_\omega = V_\omega(\omega I_k + N), \tag{6.6}$$

where N is a $k \times k$ matrix with ones on the first superdiagonal and zeros elsewhere. Then for every point $\alpha \notin \sigma(-A^*)$

$$(R_\alpha F)(\lambda) V_\omega = \mu F(\lambda) V_\omega (I_k - \mu N)^{-1}, \tag{6.7}$$

where $\mu = -(\omega + \alpha)^{-1}$.

Proof. Formula (6.6) implies that

$$(\alpha I_n + A^*) V_\omega = V_\omega((\omega + \alpha) I_k + N) = -\mu^{-1} V_\omega (I_k - \mu N)$$

and hence that

$$-(\alpha I_n + A^*)^{-1} V_\omega = \mu V_\omega (I_k - \mu N)^{-1}.$$

This leads easily to (6.7) with the help of (6.1). □

Lemma 6.4. *Let N be a $k \times k$ matrix with ones on the first superdiagonal and zeros elsewhere, and let $\mu \in \mathbb{C} \setminus \{0\}$. Then the matrix $\mu(I_k - \mu N)^{-1}$ is similar to the matrix $\mu I_k + N$. Moreover, the columns $v_1, \ldots, v_k$ of the $k \times k$ matrix $P = [v_1, \ldots, v_k]$ form a Jordan chain for the matrix $\mu(I_k - \mu N)^{-1}$, (i.e., $\mu(I_k - \mu N)^{-1} = P(\mu I_k + N)P^{-1}$) if and only if*

$$v_{k-j} = \mu^{2j}(I_k - \mu N)^{-j} N^j v_k \quad \text{for } j = 0, \ldots, k-1 \tag{6.8}$$

and

$$\text{the bottom entry of } v_k \text{ is nonzero.} \tag{6.9}$$

Proof. It is readily checked that

$$\mu(I_k - \mu N)^{-1} - \mu I_k = \mu^2 N + \cdots + \mu^k N^{k-1}$$

is a (strictly upper triangular) $k \times k$ matrix of rank $k-1$ if $\mu \neq 0$. Hence as μ is the only eigenvalue of both $\mu(I_k - \mu N)^{-1}$ and of $\mu I_k + N$ and has algebraic multiplicity k and the same geometric multiplicity 1 in both matrices, they must be similar. Now, if P is an invertible $k \times k$ matrix such that

$$\mu(I_k - \mu N)^{-1} P = P(\mu I_k + N)$$

(i.e., if the columns of P form a Jordan chain of $\mu(I_k - \mu N)^{-1}$), then

$$PN = (\mu(I_k - \mu N)^{-1} - \mu I_k)P = \mu^2(I_k - \mu N)^{-1} NP. \tag{6.10}$$

Therefore, upon matching columns of both sides of (6.10), we have

$$v_{j-1} = \mu^2(I_k - \mu N)^{-1} N v_j, \quad \text{for } j = 2, \ldots, k.$$

Formula (6.8) is obtained by iterating the last formula starting with $j = k$:

$$v_{k-1} = \mu^2(I_k - \mu N)^{-1} N v_k, \quad v_{k-2} = \mu^2(I_k - \mu N)^{-1} N v_{k-1}, \text{ etc.}$$

Moreover, the corresponding matrix P will be upper triangular and its diagonal entries p_{jj}, $j = 1, \ldots, k$, satisfy the recursion

$$p_{j-1,j-1} = \mu^2 p_{jj}, \quad j = 2, \ldots, k.$$

Therefore, P is invertible if and only if $p_{kk} \neq 0$.

The opposite direction follows by reversing the arguments. □

We are now ready to establish the following result:

Theorem 6.5. *Let (A, B) be a stabilizable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$, and let $X = R_{st}(A, B)$, and let $v_1, \ldots, v_k \in \mathscr{R}(X)$. Then $\{v_1, \ldots, v_k\}$ is a Jordan chain of A^* corresponding to $\lambda \in \Pi_+$ if and only if for every invertible $k \times k$ matrix P that satisfies*

$$\mu(I_k - \mu N)^{-1} = P(\mu I_k + N)P^{-1} \tag{6.11}$$

the columns of FVP form a Jordan chain of R_α ($\alpha \in \mathbb{C} \setminus \sigma(-A^*)$) corresponding to $\mu = -(\lambda + \alpha)^{-1}$, where $V = [v_1, \ldots, v_k]$.

Proof. Assume first that the columns of the matrix $V = [v_1 \dots v_k]$ form a Jordan chain of A^*, i.e., $A^*V = V(\lambda I_k + N)$, and that P is a $k \times k$ invertible matrix that satisfies (6.11). Then, by Lemma 6.3,

$$(R_\alpha F)V = FV\mu(I - \mu N)^{-1} = FVP(\mu I_k + N)P^{-1}$$

and so

$$R_\alpha FVP = FVP(\mu I_k + N).$$

Since (A, B) is stabilizable, , the columns of FVP are linearly independent in $\mathscr{M}_X$ and so form a Jordan chain for R_α.

Conversely, suppose that $R_\alpha FVP = FVP(\mu I_k + N)$, where P satisfies (6.11), and the columns of V are linearly independent. Then, since R_α is a bijective map of $\mathscr{M}_X$ onto itself, $\mu \neq 0$ and

$$R_\alpha FV = FVP(\mu I_k + N)P^{-1} = FV\mu(I_k - \mu N)^{-1}.$$

Therefore, since $R_\alpha FV = -F(\alpha I_n + A^*)^{-1}V$ by (6.1), it follows that

$$-(\alpha I_n + A^*)V = \frac{V}{\mu}(I_k - \mu N) = -(\alpha + \lambda)V(I_k + (\alpha + \lambda)^{-1}N), \tag{6.12}$$

with $\mu = -(\lambda + \alpha)^{-1}$. But this in turn leads easily to the identity

$$A^*V = V(\lambda I_k + N). \tag{6.13}$$

Moreover, since the columns of *FVP* are linearly independent in $\mathscr{M}_X$,, the vectors $v_1, \dots, v_k$ are linearly independent, and, by (6.13), they form a Jordan chain of A^*, corresponding to λ. Since these vectors are in $\mathscr{R}(X)$, Lemma 6.2 insures that $\lambda \in \Pi_+$. $\square$

Remark 6.6. Let (A, B) be a stabilizable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$, and let $X = R_{st}(A, B)$. Then $(zI_n + A^*)^{-1}X$ is analytic in $\overline{\Pi_+}$.

Proof. It suffices to show that the coefficients C_j in the Laurent expansion $(zI_n + A^*)^{-1}X = \sum_{j=-\infty}^{\infty}(z - z_0)^j C_j$ about each point $z_0 \in \sigma(-A^*) \cap \Pi_+$ are all equal to zero when $j < 0$. Take $r > 0$ small enough so that $\{|z - z_0| \leqslant r\} \subset \Pi_+$ and $\{|z - z_0| \leqslant r\} \cap \sigma(-A^*) = \{z_0\}$. Then, as

$$(zI_n + A^*)^{-1}X = X(zI_n - \widehat{A})^{-1} \quad \text{for } z \notin \sigma(-A^*) \cup \sigma(\widehat{A}),$$

the formula for the *j*'th coefficient can be reexpressed as

$$\begin{aligned} C_j &= \frac{1}{2\pi \mathrm{i}} \int_{|z-z_0|=r} (z - z_0)^{-(1+j)}(zI_n + A^*)^{-1}X\,\mathrm{d}z \\ &= \frac{1}{2\pi \mathrm{i}} \int_{|z-z_0|=r} (z - z_0)^{(-1+j)}X(zI_n - \widehat{A})^{-1}\,\mathrm{d}z = 0, \end{aligned}$$

since $\sigma(\widehat{A}) \subset \Pi_-$. $\square$

It can be shown (see, e.g., [6]) that if in addition $X \geqslant 0$, then

$$\begin{aligned}\mathscr{M}_X &= H_2^m(\Pi_+) \boxminus \Theta H_2^m(\Pi_+) \\ &:= \{f \in H_2^m(\Pi_+) : \langle Jf, \Theta g\rangle_{H_2^m(\Pi_+)} = 0,\ \forall\, g \in H_2^m(\Pi_+)\}.\end{aligned}$$

Proof of Theorem 1.3. The equivalence of (1), (2), and (3) follows from the discussion in Section 4.1 and from Theorem 4.6. Definitions 4.2, 4.7, Lemma 3.3 and Corollary 2.17 imply the equivalence between (3) and (4). Finally, since (1) $\Rightarrow$ (5) is immediate from Theorem 1.2, it remains only to check that (5) $\Rightarrow$ (1). If $(M_X, \langle\,\rangle)$ is a RKHS of the de Branges type, then by Theorem 1.2 (adapted to the present notation), the matrix X is a positive semidefinite solution of (1.2), and $\Theta(z) = \Theta_X(z)$ is uniquely determined by the formula (1.10) up to J-unitary constant multiplier on the right. Thus, as

$$\begin{aligned}\det\{\Theta_X(z)\} &= \det\{I_n - (zI_n - \widehat{A})^{-1} BJB^* X\} \\ &= \det\{(zI_n - \widehat{A})^{-1})(zI_n - A)\} \\ &= \frac{\det(zI_n - A)}{\det(zI_n - \widehat{A})},\end{aligned}$$

it is readily seen that $\widehat{A}$ is stable under (1.12). $\square$

7. Null chains and Jordan chains

Assume that (A, B) is a stabilizable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$ and that $X = R_{st}(A, B)$. Then Lemma 6.2 implies that $\dim \mathscr{M}_X = \dim \mathscr{R}(X) = \#(\sigma(A) \cap \Pi_+)$ (counting multiplicities). Let $\Theta(z)$ be a minimal pole conjugator of the pair (A, B) w.r.t. Π_+. By the minimality property of $\Theta(z)$, we have also $\dim \mathscr{M}_X = \#(\text{zeros of})\ \Theta(z))$. It turns out that the connection between the two is even stronger.

In order to analyze this connection we shall need the following lemmas:

Lemma 7.1 [15]. *Let $G(z)$ be a $p \times m$ rmvf, and let $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_r$ be the indices in the local SM-form of $G(z)$ at z_0. Then for each $1 \leqslant \ell \leqslant r$*

$$k_1 + k_2 + \cdots + k_\ell = \min_{\substack{1 \leqslant i_1 < \cdots < i_\ell \leqslant p,\\ 1 \leqslant j_1 < \cdots < j_\ell \leqslant m,}} \left\{ n\left(A^{(i_1,\ldots,i_\ell)}_{(j_1,\ldots,j_\ell)}, z_0 \right) \right\} \tag{7.1}$$

where $A^{(i_1,\ldots,i_\ell)}_{(j_1,\ldots,j_\ell)}$ is the minor of order ℓ of $G(z)$, defined by the i_1'th, $\ldots$, i_ℓ'th rows and j_1'th, $\ldots$, j_ℓ'th columns of $G(z)$.

We remark that the lemma is obvious when

$$G(z) = \begin{bmatrix} \operatorname{diag}\{(z-z_0)^{k_1}, \ldots, (z-z_0)^{k_r}\} & O_{r\times(m-r)} \\ O_{(p-r)\times r} & O_{(p-r)\times(m-r)} \end{bmatrix}$$

The formula for the general case is obtained by showing that the operations which have to be performed to achieve the local SM-form do not affect $n\left(A^{(i_1,\ldots,i_\ell)}_{(j_1,\ldots,j_\ell)}, z_0\right)$, for every choice of $1 \leqslant i_1 < \cdots < i_\ell \leqslant p$; $1 \leqslant j_1 < \cdots < j_\ell \leqslant m$, and hence do not affect the right hand side of (7.1).

Observe that (7.1) determines $k_1, k_2, \ldots, k_r$ and is useful, for example, to calculate the local SM-form of $zI - A$.

Lemma 7.2. *Let $G(z) = \operatorname{diag}\{G_1(z), G_2(z), \ldots, G_k(z)\}$, where $G_1(z), \ldots, G_k(z)$ are rmvf's. Then the local SM-form of $G(z)$ at z_0 is obtained by arranging the indices of the local SM-forms of $G_1(z), G_2(z), \ldots, G_k(z)$ at z_0 in non-decreasing order.*

Example 7.3. In the setting of Lemma 7.2, let $k = 2$ and let the local SM-forms of $G_1(z)$ and $G_2(z)$ at z_0 be $\operatorname{diag}\{(z-z_0)^{-3}, (z-z_0)^{-2}, 1\}$ and $\operatorname{diag}\{(z-z_0)^{-3}, z-z_0, (z-z_0)^3\}$, respectively. Then the local SM-form of $G(z)$ at z_0 is $\operatorname{diag}\{(z-z_0)^{-3}, (z-z_0)^{-3}, (z-z_0)^{-2}, 1, (z-z_0), (z-z_0)^3\}$.

Lemma 7.4. *Let (A, B) be a controllable (stabilizable) pair and let $\operatorname{rank}(B) = r$. Let $\lambda \in \mathbb{C}$ ($\lambda \in \overline{\Pi_+}$) be an eigenvalue of A with ℓ corresponding Jordan chains of sizes $k_1 \geqslant \cdots \geqslant k_\ell$. Then $\ell \leqslant r$, and the local SM-forms of $(zI_n - A)^{-1}$ and of $(zI_n - A)^{-1}B$ at λ are*

$$\operatorname{diag}\{(z-\lambda)^{-k_1}, \ldots, (z-\lambda)^{-k_\ell}, 1, \ldots, 1\}$$

and

$$\begin{bmatrix} \operatorname{diag}\{(z-\lambda)^{-k_1}, \ldots, (z-\lambda)^{-k_\ell}; (z-\lambda)^{t_{\ell+1}}, \ldots, (z-\lambda)^{t_r}\} & O_{r\times(m-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(m-r)} \end{bmatrix}$$

respectively, where $t_{\ell+1}, \ldots, t_r$ are non-negative integers.

Proof. By moving to a left similar pair we can assume that A is in Jordan form.

Let $\lambda \in \Pi_+$ be an eigenvalue of A, with ℓ corresponding Jordan chains of sizes $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_\ell$. By Lemmas 7.1 and 7.2, the local SM-form of $zI_n - A$ at λ is $\operatorname{diag}\{1, \ldots, 1, (z-\lambda)^{k_\ell}, \ldots, (z-\lambda)^{k_1}\}$. Thus the local SM-form of $(zI_n - A)^{-1}$ at λ is $\operatorname{diag}\{(z-\lambda)^{-k_1}, (z-\lambda)^{-k_2}, \ldots, (z-\lambda)^{-k_\ell}, 1, \ldots, 1\}$. Let $t_1 \leqslant \cdots \leqslant t_r$ denote the indices of the local SM-form of $(zI_n - A)^{-1}B$ at λ. By (a) of Lemma 2.6 (by (b) of Lemma 2.6), $\ell \leqslant r$. Let $1 \leqslant s \leqslant r$ be the largest number such that $t_s < 0$. Thus,

$$\begin{aligned} &E_1(z)\operatorname{diag}\{(z-\lambda)^{-k_1}, \ldots, (z-\lambda)^{-k_\ell}, 1, \ldots, 1\}F_1(z)B \\ &\quad = E_2(z)\begin{bmatrix} \operatorname{diag}\{(z-\lambda)^{t_1}, \ldots, (z-\lambda)^{t_s}, \ldots, (z-\lambda)^{t_r}\} & O_{r\times(m-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(m-r)} \end{bmatrix}F_2(z), \end{aligned}$$

where $E_1(z)$, $E_2(z)$ and $F_1(z)$ are $n \times n$ rmvf's, $F_2(z)$ is an $m \times m$ rmvf and all four of them are analytic and invertible at λ. Since B is a constant matrix,

$$s \leqslant \ell. \tag{7.2}$$

By Lemma 7.1 and the Binet–Cauchy formula, we deduce that for $1 \leqslant i \leqslant \ell$

$$-k_i \leqslant t_i. \tag{7.3}$$

Since both $(zI_n - A)^{-1}$ and $(zI_n - A)^{-1}B$ are minimal realizations, Kalman's Theorem (Theorem 2.15) guarantees that

$$-k_1 - \cdots - k_\ell = t_1 + \cdots + t_s. \tag{7.4}$$

Formulas (7.2), (7.3) and (7.4) imply that $s = \ell$ and $-k_i = t_i$ for $1 \leqslant i \leqslant \ell$. Thus, the local SM-form of $(zI_n - A)^{-1}B$ at λ is

$$\begin{bmatrix} \operatorname{diag}\{(z-\lambda)^{-k_1}, \ldots, (z-\lambda)^{-k_\ell}; (z-\lambda)^{t_{\ell+1}}, \ldots, (z-\lambda)^{t_r}\} & O_{r\times(m-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(m-r)} \end{bmatrix},$$

where $0 \leqslant t_{\ell+1} \leqslant \cdots \leqslant t_r$. $\square$

Let (A, B) be a stabilizable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$, let $\lambda \in \Pi_+$ be an eigenvalue of A with ℓ Jordan chains of sizes $k_1 \geqslant k_2 \geqslant \cdots k_\ell$ and let $\Theta(z)$ be a minimal pole conjugator of the pair (A, B) w.r.t. Π_+. By the minimality of $\Theta(z)$ and Lemma 7.4, we deduce that the local SM-form of $\Theta(z)$ at $z = \lambda$ is

$$\operatorname{diag}\{(z-\lambda)^{k_\ell}, \ldots, (z-\lambda)^{k_1}, 1, \ldots, 1\}. \tag{7.5}$$

The analysis in [3] guarantees the existence of a set $f_1(z), \ldots, f_\ell(z)$ of $1 \times m$ rmvf's that are analytic and nonzero at λ such that the matrix (A_λ, B_λ) that will be defined below is a left null pair of $\Theta(z)$ at λ:

$$A_\lambda = \operatorname{diag}\{A_1^\lambda, \ldots, A_\ell^\lambda\}, \quad B_\lambda = \begin{bmatrix} B_1^\lambda \\ \vdots \\ B_\ell^\lambda \end{bmatrix}$$

where, for $1 \leqslant j \leqslant \ell$,

$$A_j^\lambda = \lambda I_{k_j} + N_{k_j}, \quad \text{and} \quad B_j^\lambda = \begin{bmatrix} f_j(\lambda) \\ \vdots \\ f_j^{(k_j-1)}(\lambda)/(k_j-1)! \end{bmatrix}.$$

The ordered set of vectors $\left\{f_j(\lambda), \ldots, \frac{f_j^{(k_1-1)}(\lambda)}{(k_j-1)!}\right\}$ is called a left null chain of $\Theta(z)$ at λ of order k_j. The ordered set of chains $\left\{f_1(\lambda), \ldots, \frac{f_1^{(k_1-1)}(\lambda)}{(k_1-1)!}\right\}; \ldots;$ $\left\{f_\ell(\lambda), \ldots, \frac{f_\ell^{(k_\ell-1)}(\lambda)}{(k_\ell-1)!}\right\}$ is called a canonical set of left null chains of $\Theta(z)$ at λ.

We thus have the following conclusions:

Theorem 7.5. *Let (A, B) be a stabilizable pair with $\sigma(A) \cap \mathrm{i}\mathbb{R} = \emptyset$. Let $\Theta(z)$ be a minimal pole conjugator of (A, B) w.r.t. Π_+ and let $\lambda \in \Pi_+ \cap \sigma(A)$. Then the number and sizes of Jordan chains of A corresponding to λ is the same as the number and sizes of chains in (every) canonical set of left null chains of $\Theta(z)$ at λ.*

More generally, if $\Omega \subset \mathbb{C}$ and $\lambda_1, \ldots, \lambda_t$ are all the zeros of $\Theta(z)$ in Ω, then the pair (A, B), where $A = \mathrm{diag}\{A_{\lambda_1}, \ldots, A_{\lambda_t}\}$, $B = \begin{bmatrix} B_{\lambda_1} \\ \vdots \\ B_{\lambda_t} \end{bmatrix}$ is called a left null pair of $\Theta(z)$ w.r.t. Ω as is (TAT^{-1}, TB) for any invertible matrix T. It turns out that a left null pair is always controllable.

The original definitions of left null chains and left null pairs that were formulated in [3] did not make use of the local SM-form. However, as they noted, and, as is indicated above, the number and sizes of the chains in any such set is determined by the local SM-form, i.e., ℓ chains of decreasing sizes $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_\ell$. The notion of a left null pair that was introduced earlier in Definition 4.7 is of course equivalent to the definition given in [3].

Canonical sets of left pole chains and left pole pairs may be defined analogously. However, for the goals of this paper, it turned out to be more convenient to introduce left pole pairs via Definition 2.14, which is based on the analysis in Sections 3.3 and 3.4 of [3].

The interested reader is referred to [3] for a detailed exposition of null structure theory and pole structure theory. We also remark that the notions in [3] are first discussed 'from the right', i.e., a canonical set of right null (pole) chains is introduced in order to construct right null (pole) pairs. The definitions of left null (pole) chains and left null (pole) pairs may be introduced in a completely analogous way, or by applying transposes, as noted in [3]. In this paper we were interested in the 'left case', since the conjugator Θ multiplies $G(z)$ on the right.

Acknowledgments

This research was partially supported by the German-Israeli Foundation for Scientific Research and Development under a G.I.F. grant and by the Israel Science Foundation under grant 300/02. HD thanks Renee and Jay Weiss for endowing the Chair that supports his research.

References

[1] D. Alpay, I. Gohberg, Unitary rational matrix functions, in: I. Gohberg (Ed.), Topics in Interpolation Theory of Rational Matrix-Valued Functions, Operator Theory: Advances Applications, vol. 33, Birkhäuser, Basel, 1988, pp. 175–222.

[2] H. Bart, I. Gohberg, M.A. Kaashoek, Minimal Factorization of Matrix aqd Operator Functions, Birkhäuser, Basel, 1979.
[3] J.A. Ball, I. Gohberg, L. Rodman, Interpolation of Rational Matrix Functions, Birkhäuser, Basel, 1990.
[4] P. Dewilde, H. Dym, Lossless chain scattering matrices and optimum linear prediction: the vector case, Circuit Theory Appl. 9 (1981) 135–175.
[5] P. Van Dooren, Rational and polynomial matrix factorizations via recursive pole-zero cancellation, Linear Algebra Appl. 137/138 (1990) 663–697.
[6] H. Dym, J Contractive Matrix Functions, Reproducing Kernel Hilbert Spaces and Interpolation, CBMS Reg. Conf. Ser. in Math., vol. 71, American Mathematical Society, Providence, RI, 1989.
[7] H. Dym, On Riccati equations and reproducing kernel spaces, in: A. Dijksma, M.A. Kaashoek, A.C.M. Ran (Eds.), Recent Advances in Operator Theory, The Israel Gohberg Anniversary Volume, Operator Theory: Advances and Applications, vol. 124, Birkhäuser, Basel, 2001, pp. 189–215.
[8] H. Dym, Riccati equations and interpolation, in: V. Olshevsky (Ed.), Fast Algorithms, Contemporary Mathematics, vol. 323, American Mathematical Society, Providence, RI, 2003, pp. 361–391.
[9] H. Dym, Linear fractional transformations, Riccati equations and bitangential interpolation, revisited, in: D. Alpay (Ed.), Reproducing Kernel Spaces and Applications, Operator Theory: Advances Applications, vol. 143, Birkhäuser, Basel, 2003, pp. 171–212.
[10] H. Dym, S. Nevo, Zero cancellation, Linear Algebra Appl., in press (doi:10.1016/j.laa.2005. 02.015).
[11] C. Foias, A.E. Frazho, The Commutant Lifting Approach to Interpolation Problems, Birkhäuser, Basel, 1990.
[12] I. Gohberg, M.A. Kaashoek, A. Ran, Partial pole and zero displacements by cascade connection, SIAM J. Matrix Anal. Appl. 10 (1989) 316–325.
[13] M. Green, H_∞ controller synthesis by J-lossless coprime factorization, SIAM J. Control Opt. 30 (1992) 522–547.
[14] M. Green, K. Glover, D.J.N. Limebeer, J.C. Doyle, A J-spectral factorization approach to H_∞ control, SIAM J. Control Opt. 28 (1990) 1350–1377.
[15] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980.
[16] R.E. Kalman, Irreducible realizations and the degree of a rational matrix, J. Soc. Indust. Appl. Math. 13 (1965) 520–544.
[17] V.E. Katsnelson, A.Ya. Kheifets, P.M. Yuditskii, An abstract interpolation problem and the extension theory of isometric operators, in: H. Dym, B. Fritsche, V.E. Katsnelson, B. Kirstein (Eds.), Topics in Interpolation Theory, Operator Theory: Advances and Applications, vol. 95, Birkhäuser, Basel, 1997, pp. 283–298.
[18] H. Kimura, Chain Scattering Approach to H_∞-Control, Birkhäuser, Boston, 1997.
[19] M. Rakowsky, Zero-pole interpolation of nonregular rational matrix functions, Ph.D. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1989.